

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 101788,847A  
Source: JFwo  
Date Processed by STIC: 12/8/06

***ENTERED***

**CRF Errors Edited by the STIC Systems Branch**

Serial Number: 1017881847A

CRF Edit Date: 12/18/06  
Edited by: Xe

\_\_\_\_\_ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

\_\_\_\_\_ Corrected the SEQ ID NO. Sequence numbers edited were:  
\_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:  
\_\_\_\_\_  
\_\_\_\_\_

✓ Deleted: ✓ invalid beginning/end-of-file text ; \_\_\_\_ page numbers

\_\_\_\_\_ Inserted mandatory headings/numeric identifiers, specifically:  
\_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_ Moved responses to same line as heading/numeric identifier, specifically:  
\_\_\_\_\_  
\_\_\_\_\_

\_\_\_\_\_ Other:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_



IFWO

**RAW SEQUENCE LISTING** DATE: 12/13/2006  
**PATENT APPLICATION:** US/10/788,847A TIME: 16:15:05

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\12132006\J788847A.raw

3 <110> APPLICANT: Nakamura, Yusuke  
4 Furukawa, Yoichi  
6 <120> TITLE OF INVENTION: Gene and Protein Relating to Hepatocellular  
Carcinoma and Methods  
7 of Use Thereof  
9 <130> FILE REFERENCE: 25371-021 CIP  
11 <140> CURRENT APPLICATION NUMBER: US 10/788,847A  
12 <141> CURRENT FILING DATE: 2004-02-27  
14 <150> PRIOR APPLICATION NUMBER: PCT/JP02/09876  
15 <151> PRIOR FILING DATE: 2002-09-25  
17 <150> PRIOR APPLICATION NUMBER: US 60/324,261  
18 <151> PRIOR FILING DATE: 2001-09-25  
20 <150> PRIOR APPLICATION NUMBER: US 60/391,666  
21 <151> PRIOR FILING DATE: 2002-06-26  
23 <150> PRIOR APPLICATION NUMBER: CASN 2,399,569  
24 <151> PRIOR FILING DATE: 2002-08-23  
26 <150> PRIOR APPLICATION NUMBER: 60/450,644  
27 <151> PRIOR FILING DATE: 2003-02-28  
29 <160> NUMBER OF SEQ ID NOS: 83  
31 <170> SOFTWARE: PatentIn version 3.2  
33 <210> SEQ ID NO: 1  
34 <211> LENGTH: 1622  
35 <212> TYPE: DNA  
36 <213> ORGANISM: Homo sapiens  
39 <220> FEATURE:  
40 <221> NAME/KEY: CDS  
41 <222> LOCATION: (96)..(1382)  
43 <400> SEQUENCE: 1  
44 gtgcgcgcag ggccgcaggcg cgccgggtccc ggcagccccgt gagacgccccg ctgctggacg 60  
46 cggtagccg tctgagggtgc cggagctgcg ggagg atg gag ccg ctg aag gtg 113  
47 Met Glu Pro Leu Lys Val  
48 1 5  
50 gaa aag ttc gca acc gcc aac agg gga aac ggg ctg cgc gcc gtg acc 161  
51 Glu Lys Phe Ala Thr Ala Asn Arg Gly Asn Gly Leu Arg Ala Val Thr  
52 10 15 20  
54 ccg ctg cgc ccc gga gag cta ctc ttc cgc tcg gat ccc ttg gcg tac 209  
55 Pro Leu Arg Pro Gly Glu Leu Leu Phe Arg Ser Asp Pro Leu Ala Tyr  
56 25 30 35  
58 acg gtg tgc aag ggg agt cgt ggc gtc gtc gac cgc tgc ctt ctc 257  
59 Thr Val Cys Lys Gly Ser Arg Gly Val Val Cys Asp Arg Cys Leu Leu  
60 40 45 50  
62 ggg aag gaa aag ctg atg cga tgc tct cag tgc cgc gtc gcc aaa tac 305  
63 Gly Lys Glu Lys Leu Met Arg Cys Ser Gln Cys Arg Val Ala Lys Tyr  
64 55 60 65 70

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/10/788,847A**

**DATE: 12/13/2006**  
**TIME: 16:15:05**

**Input Set : A:\pto.kd.txt**  
**Output Set: N:\CRF4\12132006\J788847A.raw**

66 tgt agt gct aag tgt cag aaa aaa gct tgg cca gac cac aag cg	353
67 Cys Ser Ala Lys Cys Gln Lys Lys Ala Trp Pro Asp His Lys Arg Glu	
68 75 80 85	
70 tgc aaa tgc ctt aaa agc tgc aaa ccc aga tat cct cca gac tcc gtt	401
71 Cys Lys Cys Leu Lys Ser Cys Lys Pro Arg Tyr Pro Pro Asp Ser Val	
72 90 95 100	
74 cga ctt ctt ggc aga gtt gtc ttc aaa ctt atg gat gga gca cct tca	449
75 Arg Leu Leu Gly Arg Val Val Phe Lys Leu Met Asp Gly Ala Pro Ser	
76 105 110 115	
78 gaa tca gag aag ctt tac tca ttt tat gat ctg gag tca aat att aac	497
79 Glu Ser Glu Lys Leu Tyr Ser Phe Tyr Asp Leu Glu Ser Asn Ile Asn	
80 120 125 130	
82 aaa ctg act gaa gat aag aaa gag ggc ctc agg caa ctc gta atg aca	545
83 Lys Leu Thr Glu Asp Lys Lys Glu Gly Leu Arg Gln Leu Val Met Thr	
84 135 140 145 150	
86 ttt caa cat ttc atg aga gaa ata cag gat gcc tct cag ctg cca	593
87 Phe Gln His Phe Met Arg Glu Glu Ile Gln Asp Ala Ser Gln Leu Pro	
88 155 160 165	
90 cct gcc ttt gac ctt ttt gaa gcc ttt gca aaa gtg atc tgc aac tct	641
91 Pro Ala Phe Asp Leu Phe Glu Ala Phe Ala Lys Val Ile Cys Asn Ser	
92 170 175 180	
94 ttc acc atc tgt aat gcg gag atg cag gaa gtt ggt gtt ggc cta tat	689
95 Phe Thr Ile Cys Asn Ala Glu Met Gln Glu Val Gly Val Gly Leu Tyr	
96 185 190 195	
98 ccc agt atc tct ttg ctc aat cac agc tgt gac ccc aac tgt tcg att	737
99 Pro Ser Ile Ser Leu Leu Asn His Ser Cys Asp Pro Asn Cys Ser Ile	
100 200 205 210	
102 gtg ttc aat ggg ccc cac ctc tta ctg cga gca gtc cga gac atc gag	785
103 Val Phe Asn Gly Pro His Leu Leu Leu Arg Ala Val Arg Asp Ile Glu	
104 215 220 225 230	
106 gtg gga gag gag ctc acc atc tgc tac ctg gat atg ctg atg acc agt	833
107 Val Gly Glu Glu Leu Thr Ile Cys Tyr Leu Asp Met Leu Met Thr Ser	
108 235 240 245	
110 gag gag cgc cgg aag cag ctg agg gac cag tac tgc ttt gaa tgt gac	881
111 Glu Glu Arg Arg Lys Gln Leu Arg Asp Gln Tyr Cys Phe Glu Cys Asp	
112 250 255 260	
114 tgt ttc cgt tgc caa acc cag gac aag gat gct gat atg cta act ggt	929
115 Cys Phe Arg Cys Gln Thr Gln Asp Lys Asp Ala Asp Met Leu Thr Gly	
116 265 270 275	
118 gat gag caa gta tgg aag gaa gtt caa gaa tcc ctg aaa aaa att gaa	977
119 Asp Glu Gln Val Trp Lys Glu Val Gln Glu Ser Leu Lys Lys Ile Glu	
120 280 285 290	
122 gaa ctg aag gca cac tgg aag tgg gag cag gtt ctg gcc atg tgc cag	1025
123 Glu Leu Lys Ala His Trp Lys Trp Glu Gln Val Leu Ala Met Cys Gln	
124 295 300 305 310	
126 gcg atc ata agc agc aat tct gaa cgg ctt ccc gat atc aac atc tac	1073
127 Ala Ile Ile Ser Ser Asn Ser Glu Arg Leu Pro Asp Ile Asn Ile Tyr	
128 315 320 325	
130 cag ctg aag gtg ctc gac tgc gcc atg gat gcc tgc atc aac ctc ggc	1121

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/788,847A

DATE: 12/13/2006

TIME: 16:15:05

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\12132006\J788847A.raw

131 Gln Leu Lys Val Leu Asp Cys Ala Met Asp Ala Cys Ile Asn Leu Gly  
 132 330 335 340  
 134 ctg ttg gag gaa gcc ttg ttc tat ggt act cgg acc atg gag cca tac 1169  
 135 Leu Leu Glu Glu Ala Leu Phe Tyr Gly Thr Arg Thr Met Glu Pro Tyr  
 136 345 350 355  
 138 agg att ttt ttc cca gga agc cat ccc gtc aga ggg gtt caa gtg atg 1217  
 139 Arg Ile Phe Phe Pro Gly Ser His Pro Val Arg Gly Val Gln Val Met  
 140 360 365 370  
 142 aaa gtt ggc aaa ctg cag cta cat caa ggc atg ttt ccc caa gca atg 1265  
 143 Lys Val Gly Lys Leu Gln Leu His Gln Gly Met Phe Pro Gln Ala Met  
 144 375 380 385 390  
 146 aag aat ctg aga ctg gct ttt gat att atg aga gtg aca cat ggc aga 1313  
 147 Lys Asn Leu Arg Leu Ala Phe Asp Ile Met Arg Val Thr His Gly Arg  
 148 395 400 405  
 150 gaa cac agc ctg att gaa gat ttg att cta ctt tta gaa gaa tgc gac 1361  
 151 Glu His Ser Leu Ile Glu Asp Leu Ile Leu Leu Glu Glu Cys Asp  
 152 410 415 420  
 154 gcc aac atc aga gca tcc taa gggAACGcag tcagaggaa atacggcgtg 1412  
 155 Ala Asn Ile Arg Ala Ser  
 156 425  
 158 tgtcttgtt gaatgccta ttgaggtcac acactctatg ctttgttagc tgtgtgaacc 1472  
 160 tctcttattt gaaattctgt tcctgtttt tgttagtaaa taaaggcaga catggttgc 1532  
 162 aaaccacaag aatcattagt tgtagagaag cacgattata ataaattcaa aacatttggt 1592  
 164 tgaggatgcc aaaaaaaaaaaaaaaa 1622  
 167 <210> SEQ ID NO: 2  
 168 <211> LENGTH: 428  
 169 <212> TYPE: PRT  
 170 <213> ORGANISM: Homo sapiens  
 172 <400> SEQUENCE: 2  
 174 Met Glu Pro Leu Lys Val Glu Lys Phe Ala Thr Ala Asn Arg Gly Asn  
 175 1 5 10 15  
 178 Gly Leu Arg Ala Val Thr Pro Leu Arg Pro Gly Glu Leu Leu Phe Arg  
 179 20 25 30  
 182 Ser Asp Pro Leu Ala Tyr Thr Val Cys Lys Gly Ser Arg Gly Val Val  
 183 35 40 45  
 186 Cys Asp Arg Cys Leu Leu Gly Lys Glu Lys Leu Met Arg Cys Ser Gln  
 187 50 55 60  
 190 Cys Arg Val Ala Lys Tyr Cys Ser Ala Lys Cys Gln Lys Lys Ala Trp  
 191 65 70 75 80  
 194 Pro Asp His Lys Arg Glu Cys Lys Cys Leu Lys Ser Cys Lys Pro Arg  
 195 85 90 95  
 198 Tyr Pro Pro Asp Ser Val Arg Leu Leu Gly Arg Val Val Phe Lys Leu  
 199 100 105 110  
 202 Met Asp Gly Ala Pro Ser Glu Ser Glu Lys Leu Tyr Ser Phe Tyr Asp  
 203 115 120 125  
 206 Leu Glu Ser Asn Ile Asn Lys Leu Thr Glu Asp Lys Lys Glu Gly Leu  
 207 130 135 140  
 210 Arg Gln Leu Val Met Thr Phe Gln His Phe Met Arg Glu Glu Ile Gln  
 211 145 150 155 160

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/788,847A

DATE: 12/13/2006  
TIME: 16:15:05

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\12132006\J788847A.raw

214 Asp Ala Ser Gln Leu Pro Pro Ala Phe Asp Leu Phe Glu Ala Phe Ala  
 215 165 170 175  
 218 Lys Val Ile Cys Asn Ser Phe Thr Ile Cys Asn Ala Glu Met Gln Glu  
 219 180 185 190  
 222 Val Gly Val Gly Leu Tyr Pro Ser Ile Ser Leu Leu Asn His Ser Cys  
 223 195 200 205  
 226 Asp Pro Asn Cys Ser Ile Val Phe Asn Gly Pro His Leu Leu Leu Arg  
 227 210 215 220  
 230 Ala Val Arg Asp Ile Glu Val Gly Glu Leu Thr Ile Cys Tyr Leu  
 231 225 230 235 240  
 234 Asp Met Leu Met Thr Ser Glu Glu Arg Arg Lys Gln Leu Arg Asp Gln  
 235 245 250 255  
 238 Tyr Cys Phe Glu Cys Asp Cys Phe Arg Cys Gln Thr Gln Asp Lys Asp  
 239 260 265 270  
 242 Ala Asp Met Leu Thr Gly Asp Glu Gln Val Trp Lys Glu Val Gln Glu  
 243 275 280 285  
 246 Ser Leu Lys Lys Ile Glu Glu Leu Lys Ala His Trp Lys Trp Glu Gln  
 247 290 295 300  
 250 Val Leu Ala Met Cys Gln Ala Ile Ile Ser Ser Asn Ser Glu Arg Leu  
 251 305 310 315 320  
 254 Pro Asp Ile Asn Ile Tyr Gln Leu Lys Val Leu Asp Cys Ala Met Asp  
 255 325 330 335  
 258 Ala Cys Ile Asn Leu Gly Leu Leu Glu Glu Ala Leu Phe Tyr Gly Thr  
 259 340 345 350  
 262 Arg Thr Met Glu Pro Tyr Arg Ile Phe Phe Pro Gly Ser His Pro Val  
 263 355 360 365  
 266 Arg Gly Val Gln Val Met Lys Val Gly Lys Leu Gln Leu His Gln Gly  
 267 370 375 380  
 270 Met Phe Pro Gln Ala Met Lys Asn Leu Arg Leu Ala Phe Asp Ilé Met  
 271 385 390 395 400  
 274 Arg Val Thr His Gly Arg Glu His Ser Leu Ile Glu Asp Leu Ile Leu  
 275 405 410 415  
 278 Leu Leu Glu Glu Cys Asp Ala Asn Ile Arg Ala Ser  
 279 420 425  
 282 <210> SEQ ID NO: 3  
 283 <211> LENGTH: 55  
 284 <212> TYPE: DNA  
 285 <213> ORGANISM: Artificial Sequence  
 287 <220> FEATURE:  
 288 <223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence  
 290 <400> SEQUENCE: 3  
 291 caccaaactt atggatggag caccttcaa gagaagggtgc tccatccata agttt 55  
 294 <210> SEQ ID NO: 4  
 295 <211> LENGTH: 55  
 296 <212> TYPE: DNA  
 297 <213> ORGANISM: Artificial Sequence  
 299 <220> FEATURE:  
 300 <223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence  
 302 <400> SEQUENCE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/788,847A

DATE: 12/13/2006  
TIME: 16:15:05

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\12132006\J788847A.raw

303	aaaaaaaactt atggatggag caccttctct tgaaaggtgc tccatccata agttt	55
306	<210> SEQ ID NO: 5	
307	<211> LENGTH: 55	
308	<212> TYPE: DNA	
309	<213> ORGANISM: Artificial Sequence	
311	<220> FEATURE:	
312	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
314	<400> SEQUENCE: 5	
315	caccaatcag agaagttta ctcatttcaa gagaatgagt aaagcttata tgatt	55
318	<210> SEQ ID NO: 6	
319	<211> LENGTH: 55	
320	<212> TYPE: DNA	
321	<213> ORGANISM: Artificial Sequence	
323	<220> FEATURE:	
324	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
326	<400> SEQUENCE: 6	
327	aaaaaatcag agaagttta ctcattctct tgaaatgagt aaagcttata tgatt	55
330	<210> SEQ ID NO: 7	
331	<211> LENGTH: 55	
332	<212> TYPE: DNA	
333	<213> ORGANISM: Artificial Sequence	
335	<220> FEATURE:	
336	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
338	<400> SEQUENCE: 7	
339	caccaacaaa ctgactgaag ataagttcaa gagaaggtgc tccatccata agttt	55
342	<210> SEQ ID NO: 8	
343	<211> LENGTH: 55	
344	<212> TYPE: DNA	
345	<213> ORGANISM: Artificial Sequence	
347	<220> FEATURE:	
348	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
350	<400> SEQUENCE: 8	
351	aaaaaacaaa ctgactgaag ataagtctct tgaaaggtgc tccatccata agttt	55
354	<210> SEQ ID NO: 9	
355	<211> LENGTH: 55	
356	<212> TYPE: DNA	
357	<213> ORGANISM: Artificial Sequence	
359	<220> FEATURE:	
360	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
362	<400> SEQUENCE: 9	
363	caccaactcg taatgacatt tcaacttcaa gagagttgaa atgtcattac gagtt	55
366	<210> SEQ ID NO: 10	
367	<211> LENGTH: 55	
368	<212> TYPE: DNA	
369	<213> ORGANISM: Artificial Sequence	
371	<220> FEATURE:	
372	<223> OTHER INFORMATION: An Artificially Synthesized siRNA Sequence	
374	<400> SEQUENCE: 10	
375	aaaaaactcg taatgacatt tcaactctct tgaagttgaa atgtcattac gagtt	55

RAW SEQUENCE LISTING ERROR SUMMARY                    DATE: 12/13/2006  
PATENT APPLICATION: US/10/788,847A                    TIME: 16:15:06

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\12132006\J788847A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:37; N Pos. 21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40  
Seq#:44; N Pos. 485,486,487,488,489

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:80

**VERIFICATION SUMMARY** DATE: 12/13/2006  
PATENT APPLICATION: US/10/788,847A TIME: 16:15:06

Input Set : A:\pto.kd.txt  
Output Set: N:\CRF4\12132006\J788847A.raw

L:705 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:0  
L:813 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:480

**Raw Sequence Listing before editing  
(for reference only)**



IFWO

**RAW SEQUENCE LISTING**

PATENT APPLICATION: US/10/788,847A

DATE: 12/08/2006

TIME: 14:46:51

Input Set : A:\25371-021CIP Sequence Listing.txt  
 Output Set: N:\CRF4\12082006\J788847A.raw

3 <110> APPLICANT: Nakamura, Yusuke  
 4 Furukawa, Yoichi

6 <120> TITLE OF INVENTION: Gene and Protein Relating to Hepatocellular Carcinoma and  
 Methods

7 of Use Thereof  
 9 <130> FILE REFERENCE: 25371-021 CIP  
 11 <140> CURRENT APPLICATION NUMBER: US 10/788,847A  
 12 <141> CURRENT FILING DATE: 2004-02-27  
 14 <150> PRIOR APPLICATION NUMBER: PCT/JP02/09876  
 15 <151> PRIOR FILING DATE: 2002-09-25  
 17 <150> PRIOR APPLICATION NUMBER: US 60/324,261  
 18 <151> PRIOR FILING DATE: 2001-09-25  
 20 <150> PRIOR APPLICATION NUMBER: US 60/391,666  
 21 <151> PRIOR FILING DATE: 2002-06-26  
 23 <150> PRIOR APPLICATION NUMBER: CASN 2,399,569  
 24 <151> PRIOR FILING DATE: 2002-08-23  
 26 <150> PRIOR APPLICATION NUMBER: 60/450,644  
 27 <151> PRIOR FILING DATE: 2003-02-28  
 29 <160> NUMBER OF SEQ ID NOS: 83  
 31 <170> SOFTWARE: PatentIn version 3.2

Does Not Comply  
 Corrected Diskette Needed  
 (pg.1)

**ERRORED SEQUENCES**

1623 <210> SEQ ID NO: 83  
 1624 <211> LENGTH: 30  
 1625 <212> TYPE: DNA  
 1626 <213> ORGANISM: Artificial Sequence  
 1628 <220> FEATURE:  
 1629 <223> OTHER INFORMATION: An Artificially Synthesized Oligonucleotide Sequence  
 1631 <400> SEQUENCE: 83  
 1632 gcgggaggat ggacccgctg aaggtggaaa

30

E--&gt; 1638 1

*Q deleted*

RAW SEQUENCE LISTING ERROR SUMMARY                    DATE: 12/08/2006  
PATENT APPLICATION: US/10/788,847A                    TIME: 14:46:52

Input Set : A:\25371-021CIP Sequence Listing.txt  
Output Set: N:\CRF4\12082006\J788847A.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:80

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/788,847A

DATE: 12/08/2006

TIME: 14:46:52

Input Set : A:\25371-021CIP Sequence Listing.txt  
Output Set: N:\CRF4\12082006\J788847A.raw

L:705 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37 after pos.:0  
L:813 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44 after pos.:480  
L:1638 M:254 E: No. of Bases conflict, this line has no nucleotides.